AUTOMATIC TUNE FAMILY IDENTIFICATION BY MUSICAL SEQUENCE ALIGNMENT

Patrick E. Savage
Tokyo University of the Arts, Dept. of Musicology
patsavagenz@gmail.com

Quentin D. Atkinson
Auckland University, Dept. of Psychology
q.atkinson@auckland.ac.nz

ABSTRACT

Musics, like languages and genes, evolve through a process of transmission, variation, and selection. Evolution of musical tune families has been studied qualitatively for over a century, but quantitative analysis has been hampered by an inability to objectively distinguish between musical similarities that are due to chance and those that are due to descent from a common ancestor. Here we propose an automated method to identify tune families by adapting genetic sequence alignment algorithms designed for automatic identification and alignment of protein families. We tested the effectiveness of our method against a high-quality ground-truth dataset of 26 folk tunes from four diverse tune families (two English, two Japanese) that had previously been identified and aligned manually by expert musicologists. We tested different combinations of parameters related to sequence alignment and to modeling of pitch, rhythm, and text to find the combination that best matched the ground-truth classifications. The best-performing automated model correctly grouped 100% (26/26) of the tunes in terms of overall similarity to other tunes, identifying 85% (22/26) of these tunes as forming distinct tune families. The success of our approach on a diverse, cross-cultural ground-truth dataset suggests promise for future automated reconstruction of musical evolution on a wide scale.

1. INTRODUCTION

Darwin's theory of evolution is a broad one that applies not only to biology but also to cultural forms such as language and music [21], [27]. Musicologists have long been interested in understanding how and why music evolves, particularly the three key mechanisms of 1) *transmission* between generations, 2) generation of musical *variation*, and 3) *selection* of certain variants over others [10], [21]. In some cases, historical notations, audio recordings, or other musical "fossils" allow us to document music's cultural evolution through the accumulation of minute variations over time [5], [14], [28]. More often, the process of oral transmission results in contemporaneous groups of related melodies known as "tune families" [2], careful comparison of which can be used to partially reconstruct the process of musical evolution [4]. This situation is analogous to the evolution of language families and biological species [1].

Traditionally, analysis of tune family evolution has been done by manually identifying and aligning small groups of related melodies (see Fig. 1a) and then qualitatively comparing the similarities and differences. This led to two major challenges that limited the scale of tune family research: 1) the need for an automated method of comparing large numbers of melodies; and 2) the need for an objective means of determining tune family membership.

Thanks to the rise of music information retrieval (MIR), the first challenge has been largely overcome by automated sequence alignment algorithms for identifying melodic similarity [9], [16], [23], some of which have been specifically designed for studying tune families [24-26]. However, the second challenge remains unsolved, with tune family identification considered "currently too ambitious to perform automatically" [24].

Here we propose a novel method of tune family identification inspired by molecular genetics [8]. In particular, the problem of protein family identification shares many analogies with tune family identification. Proteins are biological molecules that are constructed by joining sequences of amino acids into 3-dimensional structures that function to catalyze biochemical reactions. Meanwhile, tunes are constructed by joining sequences of notes into multidimensional melodies that function to carry song lyrics, accompany dance, etc. When attempting to identify both protein families and tune families, a major challenge is to determine whether any observed similarities are due to chance or common ancestry.

We sought to develop automated methods for identifying and aligning tune families that could be used in future large-scale studies of musical evolution throughout the world. To do this, we adapted methods designed for identifying and aligning protein families and tested their effectiveness on a cross-cultural ground-truth set of well-established tune families that had already been manually identified and aligned by expert musicologists. We then tested out different model parameters to determine which parameters are most effective at capturing the known ground-truth patterns.

2. DATA

Our ground-truth dataset consisted of 26 melodies from four contrasting tune families that had previously been

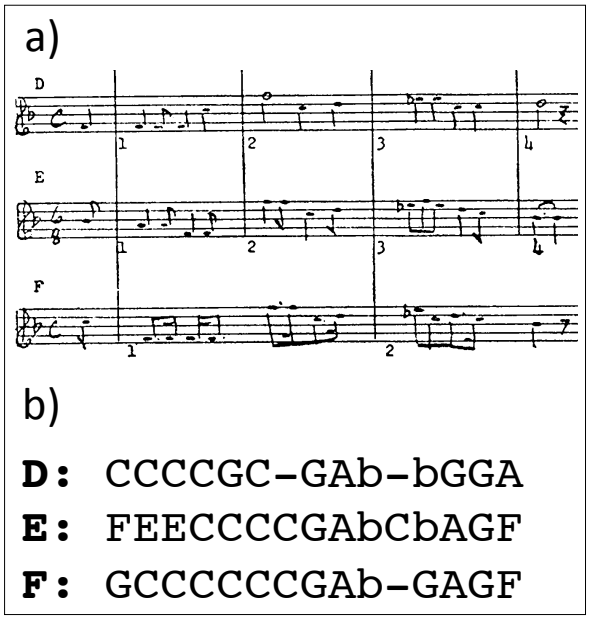

Figure 1. A sample portion of a manually aligned tune family. a) The opening phrase of three tunes manually aligned by Bayard [3] and identified as part of the tune family he labeled "Brave Donnelly". b) The same information encoded as aligned pitch-class sequences using our proposed method (see Methods and Fig. 2). Note that keys are transposed so that the tonic (originally F) is always represented as C.

identified and aligned manually by expert musicologists[1]. Two of these tune families were British-American tune families that had been chosen by Samuel Bayard (who coined the term "tune family") in order to capture "...all the problems attending a comparative tune study, and all the important features of traditional development that we constantly encounter when we try to elucidate the really extensive families of tunes." [3]. The other two were Japanese tune families chosen for similar reasons by the Japanese folksong scholars MACHIDA Kashō and TAKEUCHI Tsutomu [12]. We chose this dataset because we needed a known baseline against which to compare the effectiveness of our methods, and because we wanted our method to have cross-cultural validity that is not limited to idiosyncracies of the types of European-American folk tunes that have traditionally been studied. In addition, the first author has first-hand experience singing English and Japanese folksongs, and this dataset is also comparable to similar but larger collections of British-American and Japanese folk songs (approximately 5,000 each in [5], [18]) to which we aim to eventually apply these automated methods.

Music is much more than notes transcribed in a score. However, in order to understand tune family evolution, we need a standardized method of comparing tunes across time and space. To allow for analysis of tunes

[1] Full metadata and aligned sequences are available at http://dx.doi.org/10.6084/m9.figshare.1468015

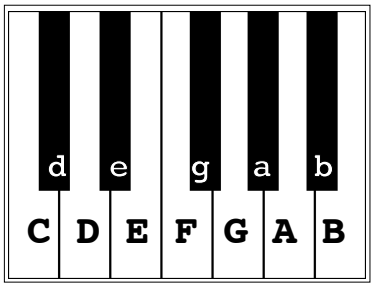

Figure 2. The most widely used "alphabet" for describing musical pitches divides an octave into 12 equally spaced semitones. Here these are visualized using the standard piano keyboard representation, with C representing the tonic.

documented before the advent of audio recording technology, this requires the use of transcriptions, although this comes at the cost of losing details about performance style (e.g., timbre, ornamentation, microtuning, microtiming). Furthermore, to allow evolutionary analysis using state-of-the-art methods from evolutionary biology, we need to further reduce the information in the score into aligned sequences. This approach was already implicit in the melodic alignment approach developed by tune family scholars, in which tunes were transposed into a common key and time signatures, phrases, and rhythms were stretched and compressed as necessary to align notes sharing similar pitches (see Fig. 1a).

Just as DNA can be modeled as a sequence constructed from an "alphabet" of 4 nucleic acids (C, G, A, or T) or a protein can be modeled as a sequence constructed from an alphabet of 20 amino acids, a melody can be modeled as a sequence constructed from an alphabet of 12 pitch classes representing the 12 notes of the chromatic scale (Fig. 2). By aligning sequences known to share common ancestry (as done manually in [3] and [12]), we can identify points on the alignment that are conserved, where a different pitch has been substituted, or where a pitch has been inserted/deleted ("indel", represented using dashes). Fig. 1b shows how this method is used to encode the manual alignment shown in Fig. 1a. This information can then be analyzed quantitatively to reconstruct a phylogenetic tree, network, or other representation of the evolutionary history of the tune family.

The intuition of early tune family scholars to emphasize alignment of pitches, rather than rhythms or global stylistic features, is supported by recent research that has demonstrated quantitatively that pitch is greatly superior to rhythm and to global stylistic features both for the purposes of tune family identification in particular and for melodic similarity in general [23], [25]. However, judicious use of rhythm and other non-pitch features may improve tune family identification [25], and we explore this using several modeling techniques.

3. METHODS

3.1 Sequence alignment parameters

Automated sequence alignment requires a number of parameters to be defined. The choice of values for these parameters depends on the nature of the data and the goals of classification. Because automated tune family identification remains largely unexplored, we don't yet know which values are most appropriate for this goal. Therefore, we tested several values for each parameter to allow for empirical comparison of which parameter values performed best. When possible, we tested values that have worked well in similar work on protein family identification and automated melodic similarity algorithms.

3.1.1 Gap penalties

The functional mechanisms of protein structure result in substitutions being much more common than indels (insertions/deletions). Thus, most amino acid alignment algorithms set a gap opening penalty (GOP) parameter to be quite high to penalize the creation of gaps in a sequence. However, when indels do occur, they often encompass not only one amino acid residue, but rather can include fairly long sections. Thus gap extension penalties (GEP) are usually set to be substantially smaller than gap opening penalties (the default values for the popular ClustalW algorithm are for GOP and GEP values of 15 and 6.66, respectively [22]).

The mechanisms of musical sequence evolution are less well known, but previous tune family research suggests that insertion/deletion (e.g., of ornamentation) is quite common and may even be more common than substitution of different pitches. Thus, it seemed desirable to examine the effect of using a range of GOP and GEP values, ranging from the combination of GOP=0.8, GEP=0.2 used to align tunes in [25], to the amino acid alignment values given above. To do this, we chose GOP values of .8, 4, 8, 12, and 16, for each of which we tested GOP:GEP ratios of both 2 and 4. Thus, the gap penalty parameters ranged from minimums of GOP=0.8, GEP=0.2 (GOP:GEP ratio=4) to maximums of GOP=16, GEP=8 (GOP:GEP ratio =2). For all gap penalty parameters we followed previous tune family research [25] in using the Needleman-Wunsch alignment algorithm [17], as implemented in the *Biostrings* package in R V3.1.1 [19].

3.1.2 Pitch

There are various possibilities for weighting pitches to accommodate different degrees of similarity beyond simple match and mismatch. Previous weighting schemes using interval consonance or interval size have shown minimal improvement over a simple match/mismatch model [25]. Here we instead explore a novel weighting scheme based on qualitative tune family research that has found that tunes will sometimes change mode (i.e., some or all scale degrees may become flattened or sharped to shift from major to minor or vice-versa [3]). To do this, we simply treated an alignment of major and minor versions of each scale degree as a match (i.e., treating lowercase letters in Fig. 2 as capitals).

3.1.3 Rhythm/text

Previous tune family research has suggested that some notes are likely to be more evolutionarily stable than others. In particular, notes that are rhythmically accented [6] or that carry text [11] are proposed to be more reliable in identifying tune families than rhythmically unaccented or non-text-carrying notes, respectively. To examine these possibilities, we contrasted the results using the full sequences with those using shorter sequences created by excluding rhythmically unaccented notes (i.e., notes not falling on the first beat of a measure) or non-text-carrying notes (e.g., notes where the vowel is held over from a previous note) from the full sequences.

3.1.4 Summary

In sum, we tested all possible combinations of the following parameters:

1) Gap opening penalty: i) .8, ii) 4, iii) 8, iv) 12 or v) 16
2) Gap opening penalty : Gap extension penalty (GOP:GEP) ratio: i) 2 or ii) 4
3) Pitch: i) including or ii) ignoring mode
4) Rhythm: i) including or ii) ignoring rhythmically unaccented notes
5) Text: i) including or ii) ignoring non-text-carrying notes

This gave a total of 5x2x2x2x2=80 parameter combinations to explore, the average values of which are reported in Table 1.

3.2 Evaluation

In order to achieve our goal of automated identification and alignment for the purpose of reconstructing tune family evolution, we need a method of quantifying how well a given alignment captures the manual judgments of experts. The goal is to maximize both the degree of match in the alignment within tune families and the degree of accuracy in separating between tune families.

3.2.1 Sequence alignment

To evaluate alignment within tune families, we need a measure of the degree to which the similarities between sequences captured by the automated alignment matched similarities captured by the manual alignments. For this, we adopted the Mantel distance matrix correlation test [13]. The Mantel *r*-value is identical to a standard Pearson correlation *r*-value, but the Mantel significance test controls for the fact that pairwise distance values in a distance matrix are not independent of one another.

We adopted the simplest method for comparing pairs of sequences, which is by calculating their percent identity (*PID)*. This is calculated based on the number of aligned pitches that are identical (*ID*) divided by the sequence length (*L)* according to the following equation:

$$PID = 100\left(\frac{ID}{\frac{L_1 + L_2}{2}}\right) \quad (1)$$

This equation uses the average length of both sequences as the denominator, as this appears to be the most consistent measure of percent identity when dealing with cases where the sequences have unequal lengths due to the insertion/deletion of large segments [15] (as occurs in our dataset).

3.2.2 Tune family identification

To evaluate separation between tune families, we need a measure of the degree to which our automated clustering into tune families matches the manual tune family classifications. This needs to take into account both true positives (tunes correctly grouped into a given tune family) and false positives (tunes incorrectly grouped into a given tune family).

A method used previously by van Kranenburg et al. [25], used the true positive rate (*tpr*) and false positive rate (*fpr*) to calculate a score *J* as follows:

$$J = \frac{tpr}{1 + fpr} \quad (2)$$

Because van Kranenburg et al. did not have a method for automatically identifying boundaries between tune families, they used a "nearest neighbor" criterion to define true positives. Thus, *J* represents the proportion of tunes whose nearest neighbor (tune with highest automatically measured similarity) is also in the same (manually identified) tune family. Here we calculate this *J* score, as well as a second *J* score that more directly tests our goal of identifying boundaries between tune families.

For this second *J* score, the criterion used to define true positives is of significant sequence similarity for each pair of tunes. Significance is assessed by a random permutation test, in which the PID value for a given pair of sequence is compared against the distribution of 100 random PID values given the same sequence lengths and compositions, as calculated by randomly reordering one of the sequences [8]. Thus, when calculating this second *J* score, bold values within the boxes in Table 2 (i.e., significant sequence similarity between pairs of tunes manually identified as belonging to the same tune family) are counted as true positives, while bold values outside of the boxes (i.e., significant sequence similarity between pairs of tunes not manually identified as belonging to the same tune family) are counted as false positives.

4. RESULTS

The average scores under the different alignment parameters are shown in Table 1, with the best-performing parameter values highlighted in bold.

4.1 Sequence alignment (within-family)

The degree to which similarities within tune families captured by the automated alignment match those captured by the manual alignments of experts are indexed by the Mantel correlation *r*-values, reported in Table 1. On average, all of the alignment parameter combinations gave similarly strong correlations ranging from *r*=.82-.85.

		Within-family	Between-family	
Automated alignment parameter	Parameter value	*r*	*J* (nearest neighbor)	*J* (significance)
GOP	.8	**0.850**	0.875	0.408
	4	0.843	0.870	0.421
	8	0.823	0.849	0.479
	12	0.833	**0.877**	**0.497**
	16	0.829	0.844	0.474
GOP:GEP ratio	**2**	0.834	0.862	**0.462**
	4	**0.837**	**0.864**	0.450
Mode	Included	**0.839**	0.841	0.445
	Ignored	0.832	**0.885**	**0.467**
Rhythmically unaccented notes	**Included**	**0.841**	**0.964**	**0.587**
	Ignored	0.830	0.762	0.325
Non-text notes	**Included**	**0.838**	**0.873**	**0.460**
	Ignored	0.833	0.853	0.452

Table 1. Mean values comparing different automated alignment parameters against manual ground-truth alignments. Best-performing values are highlighted in bold. See Methods for details.

4.2 Tune family identification (between-family)

The degree to which the automated algorithms were able to separate between tune families is indexed by the *J* scores, reported in the right-hand columns of Table 1. Using gap opening penalties of 12, ignoring mode, including non-text notes, and especially including rhythmically unaccented notes all improved tune-family identification. GOP:GEP ratios of 4 gave slightly higher *J* scores using the nearest neighbor criterion, but a ratio of 2 gave higher *J* scores using the more crucial criterion of significant pairwise sequence similarity. The specific parameter combination combining the best-performing parameter values - GOP=12, GOP:GEP ratio=2, ignoring mode, including rhythmically unaccented notes and including non-text notes - resulted in a Mantel correlation of *r*=.83 and *J* scores of *J*=1 and *J*=.64 for the nearest neighbor and significance criteria, respectively.

It was not possible to directly compare all parameters using the approach presented in [25], in part because the approach in [25] is based on sequences of pairwise melodic intervals, whereas the manual alignments that formed our ground-truth dataset were based on sequences of individual notes in relation to the tonic (i.e., tonic intervals). However, it was possible to directly compare between-family identification *J* scores using the best-performing parameter combination listed above, but using sequences of melodic intervals rather than tonic intervals. This melodic interval approach resulted in *J* scores of *J*=.88 and *J*=.33 for the nearest neighbor and significance criteria, respectively. These values were somewhat lower than the respective values using our tonic interval approach (*J*=1 and *J*=.64). However, further analyses are required to determine the degree to which incorporating

	1A	1B	1C	1D	1E	1F	2A	2B	2C	2D	2E	2F	2G	2H	2I	2J	2K	3A	3B	3C	3D	4A	4B	4C	4D	4E
1A		33	45	42	52	38																				
1B	**51**		29	37	31	28																				
1C	**59**	**47**		34	28	38																				
1D	**47**	**47**	**48**		40	32																				
1E	**62**	**54**	**43**	**45**		48																				
1F	**53**	**43**	**50**	**48**	**61**																					
2A	**41**	36	**34**	**37**	34	**36**		49	32	23	27	19	19	18	13	15	16									
2B	**35**	**38**	**44**	37	**45**	**38**	**54**		51	50	31	25	23	26	20	18	21									
2C	40	49	41	39	40	41	**47**	**57**		44	41	23	34	28	28	21	16									
2D	33	41	42	34	**33**	**35**	**45**	**54**	**61**		29	19	19	26	22	18	12									
2E	31	34	43	**39**	36	**42**	**45**	**48**	**57**	**44**		32	27	21	22	21	23									
2F	**43**	37	**41**	**36**	46	**34**	**39**	**48**	41	**45**	35		28	16	22	22	29									
2G	38	34	41	34	39	31	**34**	36	**42**	**40**	41	**55**		34	33	43	29									
2H	31	33	30	31	38	30	**37**	**45**	**41**	**43**	33	36	**47**		36	62	37									
2I	**44**	35	34	34	**45**	36	**28**	28	**42**	**35**	39	**30**	35	**46**		44	24									
2J	40	38	28	29	38	35	**26**	35	**39**	34	31	**39**	**55**	**62**	**49**		41									
2K	36	34	35	30	28	31	**31**	**41**	**46**	**45**	34	43	**45**	**48**	30	**39**										
3A	32	**51**	**36**	37	29	33	31	40	42	34	35	35	42	38	31	38	43		64	44	47					
3B	**40**	40	35	36	32	30	36	**43**	**46**	40	38	39	40	36	33	41	32	**61**		57	55					
3C	42	**42**	**38**	**35**	40	**45**	**30**	**38**	34	33	38	41	**39**	25	29	35	39	**51**	**62**		73					
3D	38	**45**	37	44	37	38	25	**36**	30	43	37	44	29	28	23	36	31	**56**	**60**	**67**						
4A	**40**	**40**	28	31	**39**	**40**	26	29	34	27	31	28	27	**31**	**40**	**32**	27	**27**	29	**29**	23		32	39	35	33
4B	32	29	33	38	39	36	**27**	29	35	28	39	30	27	24	30	30	22	**35**	32	28	38	40		43	45	44
4C	31	23	**36**	33	31	40	**26**	**31**	28	**27**	**38**	34	18	24	21	19	**25**	23	29	31	30	37	**52**		67	61
4D	35	26	27	30	33	36	**28**	26	36	28	35	31	26	25	27	22	21	26	30	21	24	41	**55**	**65**		78
4E	32	32	35	28	**39**	32	27	**30**	**40**	**32**	**41**	**33**	26	27	32	**29**	23	31	33	36	28	**42**	**62**	**56**	**62**	

Table 2. Pairwise percent identity scores among the 26 tunes. Tunes are labeled based on manual classifications by musicologists [3], [12]. Numbers correspond to the four tune families (1="Brave Donnelly", 2="Job of Journeywork", 3="Oiwake", 4= "Okesa"), letters correspond to the different variant tunes within each family. The values in the lower triangle are based on automated alignments using the best-performing parameters (GOP=12, GOP:GEP ratio=2, ignoring mode, including rhythmically unaccented notes and including non-text notes). The values in the upper triangle are based on manual alignments. Inter-tune family manual values are not shown because manual alignments were only done within tune families. Solid borders indicate automatically identified tune families in which at least three tunes are all significantly similar to one another. When these did not capture all tunes in a manually identified tune family, the manually identified boundaries are shown using dashed borders. Bold values indicates pairs whose similarities are significant at $P<.05$.

more fine-grained weighting of intervals, rhythmic information, etc. of the type used in [25] affects tune family identification using both melodic interval and tonic interval approaches.

4.3 Overall reconstruction of tune family evolution

The results of the top-performing parameter combination listed above are compared against manual classifications in Table 2 and Fig. 3. The lower triangle in Table 2 gives the raw pairwise sequence identity values, using bold text to indicate pairs of sequences whose similarities were statistically significant, while the upper diagonal gives within-family sequence identity values for the manual alignments. The mean percent identity values were somewhat higher for the automated alignments than the manual alignments within each family (45.7% vs. 33.7%, respectively). This presumably reflects the automated alignment identifying more false links, although in some cases it may also be identifying better alignments than the manual ones. Comparison with manual alignments conducted by different musicologists may help to clarify this issue in the future.

Fig. 3 summarizes the information in Table 2 visually using a NeighborNet diagram. NeighborNet is a type of phylogenetic network that is similar to a neighbor-joining tree, but allows visualization of conflicting non-tree like structure ("reticulation"). 100% of the tunes (26/26) were correctly grouped such that their nearest neighbor was a member of the same tune family, and the sub-grouping of tune family 2 also corresponded to Bayard's sub-grouping into a "long" and "short" version. However, only 85% (22/26) of these tunes were automatically grouped into a tune family using the criterion that all pairs within a family must be significantly similar to one another. Using this criterion also mis-identified the "long" and "short" versions of tune family 2 as two distinct tune

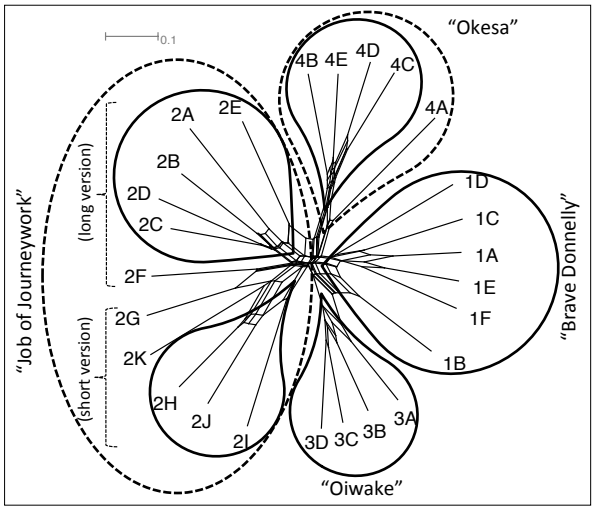

Figure 3. A NeighborNet visualization of the phylogenetic relationships among the 26 tunes automatically identified by the best-performing alignment algorithm. See Table 2 for explanation of tune labels 1A-4E and solid/dashed lines.

families. Joining families into "superfamilies" when only one or a few members have significant similarities to members of other families [8] would join the "long" and "short" versions into a superfamily, but would also join all the tune families into this superfamily.

5. DISCUSSION AND FUTURE WORK

Although previous research suggested that tune family identification was "too ambitious to perform automatically" [24], we have presented an automated approach that successfully recovers most of the key relationships within and between tune families identified manually by musicologists. Our approach adapts sequence alignment algorithms for protein family identification to successfully delineate the boundaries separating groups of melodies that share similar sequences of pitches due to descent from a common ancestor.

Our approach correctly identified three out of the four manually identified tune families, as well as both the "long version" and "short version" sub-groups of the fourth "Job of Journeywork" tune family. However, our automated approach failed to unite these sub-groups into a single tune family, instead splitting them into two tune families. The "Job of Journeywork" tune family was specifically chosen by Bayard [3] to present one of the most complicated examples of tune family evolution, including several measures that were deleted from the beginning of the "long version" and added to the end of the "short version". Hence, this type of complex evolution may require more complex algorithms and/or the incorporation of expert knowledge beyond the basic pitch sequence information encoded in the simplified model used here. However, the fact that our approach captured the relationships among the four tunes from the "Oiwake" tune family, despite the fact that this family contained both internal and terminal insertion/deletion events of substantial length, suggests that our approach is still able to capture fairly complicated patterns of musical evolution.

One area for improvement of our method is that the false positive rate is somewhat high (see Table 2). We believe that this may be due to the fact that our method is designed primarily to distinguish between chance and common ancestry, and does not do a very good job of distinguishing between common ancestry and convergent evolution. Hence, it appears likely that many of the false positives are due to stylistic similarities shared between unrelated tunes that share similar scales and motivic patterns (e.g., 1A and 2A, both Irish tunes in a diatonic major scale). Horizontal transmission and/or convergent evolution of such traits among phylogenetically unrelated groups have long been known to complicate analysis of tune family evolution [3], [7]. Horizontal transmission and convergent evolution are challenges shared with language evolution and genetic evolution, and may benefit from methods developed in these fields [1].

In the future we hope to extend our approach to larger datasets, and to incorporate more-sophisticated models of cultural evolution and sequence alignment [1], more-nuanced weighting of musical information (e.g., beyond simple match/mismatch models of pitch, rhythm, and text [24-26]), and higher-level units of musical structure and meaning. In music, as in genetics, the individual notes that make up the sequences have little meaning in themselves. The phylogenetic analysis of sequences is thus merely the starting point from which to understand how and why these sequences combine to form higher-level functional units (e.g., motives, phrases) that co-evolve with their song texts and cultural contexts of music-making as they are passed down from singer to singer through centuries of oral tradition. Using such information, we hope to not only identify previously unknown tune family relationships on a wide scale, but also to carefully reconstruct the histories and mechanisms of tune family evolution to identify general processes governing the cultural evolution of music. The general nature of our approach means that it should be applicable not only to folk music, but also to art music (e.g., European classical music [28], Japanese *gagaku* [14]) and popular music (e.g., copyright disputes [20]). Understanding the cultural evolution of music should help to identify the mechanisms that govern stability and creativity of aesthetic forms, as well as to use this knowledge to help musicians and musical cultures struggling to adapt their intangible cultural heritage to today's globalized world.

Acknowledgments: We thank H. Oota and H. Matsumae for advice on adapting genetic sequence alignment algorithms to music, and S. Brown, T. Currie, and four anonymous reviewers for comments on previous drafts of this paper. Funding support for this work was provided by a Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) scholarship to P.E.S and a Rutherford Discovery Fellowship to Q.D.A.

6. REFERENCES

[1] Q. D. Atkinson and R. D. Gray, "Curious parallels and curious connections: Phylogenetic thinking in biology and historical linguistics," *Systematic Biology*, vol. 54, no. 4, pp. 513–526, 2005.

[2] S. P. Bayard, "Prolegomena to a study of the principal melodic families of British-American folk song," *Journal of American Folklore*, vol. 63, no. 247, pp. 1–44, 1950.

[3] S. P. Bayard, "Two representative tune families of British tradition," *Midwest Folklore*, vol. 4, no. 1, pp. 13–33, 1954.

[4] C. Boiles, "Reconstruction of proto-melody," *Anuario Interamericano de Investigacion Musical*, vol. 9, pp. 45–63, 1973.

[5] B. H. Bronson, *The traditional tunes of the Child ballads: With their texts, according to the extant records of Great Britain and America [4 volumes]*. Princeton, NJ: Princeton University Press, 1959-1972.

[6] B. H. Bronson, "Toward the comparative analysis of British-American folk tunes," *Journal of American Folklore*, vol. 72, no. 284, pp. 165–191, 1959.

[7] J. R. Cowdery, "A fresh look at the concept of tune family," *Ethnomusicology*, vol. 28, no. 3, pp. 495–504, 1984.

[8] R. F. Doolittle, "Similar amino acid sequences: Chance or common ancestry?," *Science*, vol. 214, no. 4517, pp. 149–159, 1981.

[9] P. Ferraro and P. Hanna, "Optimizations of local edition for evaluating similarity between monophonic musical sequences," *Proceedings of the International Conference on Computer-Assisted Information Retrieval,* pp. 64-69, 2007.

[10] International Folk Music Council, "Resolutions: Definition of folk music," *Journal of the International Folk Music Council*, vol. 7, p. 23, 1955.

[11] A. Kaneshiro, "*Kashi onretsuhou ni yoru Oiwakebushi no hikaku* [Comparison of Oiwake melodies through lyric-note alignment]," *Minzoku Ongaku*, vol. 5, no. 1, pp. 30–36, 1990.

[12] K. Machida and T. Takeuchi, Eds., *Esashi Oiwake to Sado Okesa: Min'yo genryuukou [Folk song geneologies: Esashi Oiwake and Sado Okesa] [4 LPs]*. Kawasaki: Columbia. AL-5047/50, 1965.

[13] N. Mantel, "The detection of disease clustering and a generalized regression approach," *Cancer Research*, vol. 27, no. 2, pp. 209–220, 1967.

[14] A. Marett, "Togaku: Where have the Tang melodies gone, and where have the new melodies come from?," *Ethnomusicology*, vol. 29, no. 3, pp. 409–431, 1985.

[15] A. C. W. May, "Percent sequence identity: The need to be explicit," *Structure*, vol. 12, pp. 737–738, May 2004.

[16] M. Mongeau and D. Sankoff, "Comparison of musical sequences," *Computers and the Humanities*, vol. 24, pp. 161–175, 1990.

[17] S. B. Needleman and C. D. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *Journal of Molecular Biology*, vol. 48, pp. 443–453, 1970.

[18] NHK (Nippon Hōsō Kyōkai), Ed., *Nihon min'yō taikan [Japanese folk song anthology] [13 volumes]*. Tokyo: NHK, 1944-1994.

[19] R Development Core Team, *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2011.

[20] M. Robine, P. Hanna, P. Ferraro, and J. Allali, "Adaptation of string matching algorithms for identificaton of near-duplicate music documents," *Proceedings of the Workshop on Plagiarism Analysis, Authorship Identification, and Near-Duplicate Detection*, pp. 37–43, 2007.

[21] P. E. Savage and S. Brown, "Toward a new comparative musicology," *Analytical Approaches to World Music*, vol. 2, no. 2, pp. 148–197, 2013.

[22] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.

[23] J. Urbano, J. Lloréns, J. Morato, and S. Sánchez-cuadrado, "Melodic similarity through shape similarity," *Proceedings of the International Symposium on Computer Music Modeling and Retrieval*, pp. 338–355, 2011.

[24] P. van Kranenburg, J. Garbers, A. Volk, F. Wiering, L. Grijp, and R. C. Veltkamp, "Towards integration of MIR and folk song research," *Proceedings of the International Symposium on Music Information Retrieval*, pp. 505–508, 2007.

[25] P. van Kranenburg, A. Volk, and F. Wiering, "A comparison between global and local features for computational classification of folk song melodies," *Journal of New Music Research*, vol. 42, no. 1, pp. 1–18, 2013.

[26] P. van Kranenburg, A. Volk, F. Wiering, and R. C. Veltkamp, "Musical models for folk-song melody alignment," *Proceedings of the International Symposium on Music Information Retrieval*, pp. 507–512, 2009.

[27] A. Whiten, R. A. Hinde, C. B. Stringer, and K. N. Laland, *Culture evolves*. Oxford: Oxford University Press, 2012.

[28] H. F. Windram, T. Charlston, and C. J. Howe, "A phylogenetic analysis of Orlando Gibbons's Prelude in G," *Early Music*, vol. 42, no. 4, pp. 515–528, 2014.